

*Università degli Studi di Padova*

*Padua Research Archive - Institutional Repository*

Sampling from a Gibbs Measure with Pair Interaction by Means of PCA

*Original Citation:*

*Availability:*

This version is available at: 11577/2586044 since:

*Publisher:*

*Published version:*

DOI: 10.1007/s10955-012-0612-9

*Terms of use:*

Open Access

This article is made available under terms and conditions applicable to Open Access Guidelines, as described at <http://www.unipd.it/download/file/fid/55401> (Italian only)

(Article begins on next page)

# Sampling from a Gibbs measure with pair interaction by means of PCA

Paolo Dai Pra<sup>1</sup>

Benedetto Scoppola<sup>2</sup>

Elisabetta Scoppola<sup>3</sup>

August 13, 2012

<sup>1</sup> Dipartimento di Matematica Pura ed Applicata, University of Padova

Via Trieste 63, 35121 Padova, Italy

`daipra@math.unipd.it`

<sup>2</sup> Dipartimento di Matematica, University of Roma “Tor Vergata”

Via della Ricerca Scientifica - 00133 Roma, Italy

`scoppola@mat.uniroma2.it`

<sup>3</sup> Dipartimento di Matematica, University of Roma “Roma Tre”

Largo San Murialdo, 1 - 00146 Roma, Italy

`scoppola@mat.uniroma3.it`

## Abstract

We consider the problem of approximate sampling from the finite volume Gibbs measure with a general pair interaction. We exhibit a parallel dynamics (Probabilistic Cellular Automaton) which efficiently implements the sampling. In this dynamics the product measure that gives the new configuration in each site contains a term that tends to favour the original value of each spin. This is the main ingredient that allows one to prove that the stationary distribution of the PCA is close in total variation to the Gibbs measure. The presence of the parameter that drives the “inertial” term mentioned above gives the possibility to control the degree of parallelism of the numerical implementation of the dynamics.

# Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
1.1	Definitions . . . . .	5
1.2	Results . . . . .	7
1.3	Discussion and open problems . . . . .	9
1.4	Further remarks on the parallelization of the algorithm . . . . .	10
<b>2</b>	<b>Proofs</b>	<b>11</b>
2.1	Proof of Theorem 1.1 . . . . .	11
2.2	Föllmer's estimate . . . . .	12
2.3	Proof of Proposition 1.2 . . . . .	13
<b>3</b>	<b>The mean field Ising model</b>	<b>17</b>
3.1	Distance between $\pi^{PCA}$ and $\pi^G$ and comparison of phase diagrams . . . . .	17
3.2	Convergence to equilibrium . . . . .	20

## 1 Introduction

Probabilistic Cellular Automata (PCA) are (time-homogeneous) discrete-time Markov Chains on a product space  $S^V$ , whose transition probability  $P(d\sigma|\sigma')$  is a product measure:

$$P(\sigma|\sigma') = \prod_{i \in V} p_i(d\sigma_i|\sigma'),$$

where, for  $i \in V$  and  $\sigma' \in S^V$ ,  $p_i(d\sigma_i|\sigma')$  is a probability on  $S$ . Compared with the more familiar *sequential* dynamics, where the transition probabilities  $P(d\sigma|\sigma')$  are supported on *configuration*  $\sigma$  with  $\sigma_j = \sigma'_j$  for all but one  $j \in V$ , PCA's exhibit the following peculiar features.

- The *parallel updating* rule allows one to exploit the efficiency of *parallel computation* in the simulation of these dynamics, making them desirable Markov Chain Monte Carlo algorithms.
- PCA's give rise to well defined *infinite volume* dynamics ( $V$  infinite countable), without passing to continuous time.

The study of PCA's in the context of Equilibrium Statistical Mechanics dates back to [7, 12], where various features of the infinite-volume limit have been investigated, in particular its space-time Gibbsian nature. On the other hand, invariant measures for infinite-volume PCA's may be non-Gibbsian, as shown in [4].

In the context of Markov Chain Monte Carlo algorithms the following natural problem arises: given a probability  $\mu$  on  $S^V$ , construct a PCA whose invariant measure is  $\mu$ ; in particular, in the case  $\mu$  is a Gibbs measure for a short range interaction, one expects that the transition probabilities of the PCA can be chosen to be *local*, i.e.  $p_i(d\sigma_i|\sigma')$  depends only on  $\sigma'_j$  for  $j$  "close" to  $i$ . While Markov Chain with sequential dynamics having these features can always be constructed, the existence of a PCA with the given invariant measure  $\mu$  is not guaranteed. Counterexamples are given in [3], while [10] provides explicit conditions on  $\mu$  for the existence of a PCA *reversible* with respect to  $\mu$ .

A well understood example is that of the  $2d$  Ising model. By the results in [10] it follows that no PCA can be reversible with respect to the  $2d$  Ising model. In [12] and [1] a PCA is introduced whose invariant (reversible) measure  $\pi$  is related to the Ising model as follows: the projection of  $\pi$  to the *even* sites, i.e. those  $(i, j) \in \mathbb{Z}^2$  with  $i + j$  even, coincides with the same projection of the Ising model, and the same holds for odd sites; however, under  $\pi$ , spins at even sites are independent from spins at odd sites, unlike for the Ising model.

When the nearest neighbor interaction of the Ising model is generalized to a general pair interaction, this simple structure is lost. In this paper, following the ideas introduced in [9], we present a simple way to modify and extend the PCA in [1], and use it to sample approximately from a Gibbs measure. Given a spin configuration  $\sigma \in \{-1, 1\}^V$ , where  $V$  is a finite subset of  $\mathbb{Z}^d$ , we start with a Hamiltonian of the form

$$H(\sigma) := - \sum_{i,j} J_{ij} \sigma_i \sigma_j,$$

corresponding to the Gibbs measure

$$\pi^G(\sigma) \propto \exp[-H(\sigma)].$$

This Hamiltonian can be *lifted* to a Hamiltonian on  $(\{-1, 1\}^V)^2$ , setting

$$H(\sigma, \sigma') := - \sum_{i,j} J_{ij} \sigma_i \sigma'_j + q \sum_i (1 - \sigma_i \sigma'_i).$$

The measure

$$\mu_2(\sigma, \sigma') \propto \exp[-H(\sigma, \sigma')]$$

is such that the conditional measure

$$\mu_2(\sigma'|\sigma) = \frac{\mu_2(\sigma, \sigma')}{\sum_{\tau} \mu_2(\sigma, \tau)}$$

is a product measure, and can therefore be taken as transition probability of a PCA, which turns out to be reversible for a probability  $\pi^{PCA}$  given by the marginal of  $\mu_2$ :

$$\pi^{PCA}(\sigma) = \sum_{\tau} \mu_2(\sigma, \tau).$$

The parameter  $q$  controls the average number of spin-flips in a single step of our dynamics. This is the analogue of the self-interaction considered in [2] to study metastability in the limit of zero temperature, but our regime and our goal are completely different. The parameter  $q$  acts as the brake of the dynamics: for large values of  $q$  the dynamics is very slow, flipping few spins at each time, tending to "freeze" the system in its configuration, while a dynamics with  $q = 0$  is for instance the case of [1]. The mean density of flipped spins is of order  $\delta := e^{-2q}$ . When  $\delta$  is proportional to  $1/|V|$  we expect that the PCA behaves like the single spin flip dynamics; however we are able to prove a stronger result: in Theorem 1.2 we show that when the volume goes to infinity,  $|V| \rightarrow \infty$ , the total variation distance between  $\pi^{PCA}$  and  $\pi^G$  goes to zero when  $\lim_{|V| \rightarrow \infty} \delta^2 |V| = 0$ . In Section 2 we give the proof of this convergence for rather general two-body interactions. Given the generality of the model, convergence is proved under Dobrushin uniqueness conditions, which implies the minimal condition needed for our argument, namely a form of fast decay of correlations. For special models, such as ferromagnets in pure states (see [14]), we expect that this fast decay of correlations hold true also in the coexistence region, where Dobrushin uniqueness fails.

In other words, the control parameter  $q$  introduced in the PCA discussed in this paper, provides a "tuning" of the parallel dynamics for sampling from the Gibbs measure. This

tuning, on the one hand can be used to discuss the invariant measure of the PCA and may open new insight in this difficult field. On the other hand, from an applicative point of view, the possibility of sampling from a Gibbs measure with a parallel dynamics may be a useful tool. In subsection 1.4 we give some more details about the possibility to parallelize the algorithm in an efficient way. An application of this tool to the clique problem is discussed for instance in [9]. A general comparison between the performances of tuned PCA and single spin flip dynamics is an open problem. In the simple example of the random walk on the hypercube, corresponding to  $J_{ij} = 0$ , in [15] has been proved that the cutoff occurs at a time  $\tau^{PCA}$  smaller than the cutoff time  $\tau^G$  of the corresponding single spin flip dynamics. More precisely

$$\frac{\tau^{PCA}}{\tau^G} \sim \frac{1}{\delta|V|}.$$

In the last section of this paper we show that in the case of Curie-Weiss interaction the mixing time  $\tau_{mix}^{PCA}$  is related the cutoff time of the single spin flip dynamics, and so to its mixing time  $\tau_{mix}^G$ , by the analogue relation

$$\frac{\tau_{mix}^{PCA}}{\tau_{mix}^G} \sim \frac{1}{\delta|V|}.$$

## 1.1 Definitions

Given a finite volume  $V \subset \mathbb{Z}^d$  we consider the spin configurations  $\sigma \in \{-1, 1\}^V$  and define

$$H(\sigma) = - \sum_{i,j} J_{ij} \sigma_i \sigma_j \tag{1}$$

where  $(J_{ij})_{i,j \in \mathbb{Z}^d}$  is a given infinite *symmetric* matrix satisfying

$$\sup_i \sum_j |J_{ij}| = J < \infty \tag{2}$$

For simplicity, the sum in (1) is supposed to range over  $i, j \in V$ ; in other words, boundary conditions are free. More general boundary conditions could be treated with no difficulty. The Hamiltonian will be written equivalently

$$H(\sigma) = - \sum_{i \in V} h_i(\sigma) \sigma_i \tag{3}$$

where

$$h_i(\sigma) = \sum_j J_{ij} \sigma_j \quad (4)$$

Then we can define the standard Gibbs measure as

$$\pi^G(\sigma) = \frac{e^{-H(\sigma)}}{Z^G} \equiv \frac{w^G(\sigma)}{\sum_{\sigma} w^G(\sigma)} \quad (5)$$

where

$$Z^G = \sum_{\sigma} e^{-H(\sigma)}, \quad w^G(\sigma) = e^{-H(\sigma)} \quad (6)$$

A sampler from this Gibbs measure can be realized by defining a Markov chain defined on the state space  $\{-1, 1\}^V$  with invariant measure  $\pi^G$ , following the usual ideas of Markov Chain Monte Carlo methods. A standard algorithm is the *Gibbs sampler* which at each integer time:

- a site  $i \in V$  is randomly chosen (with uniform distribution);
- the configuration outside  $i$  is left unchanged;
- the new spin at  $i$  is sampled from the conditional measure  $\pi^G(\cdot | \sigma_{V \setminus \{i\}})$

Thus the transition matrix for the Markov chain is

$$P_{\sigma, \tau}^G = \begin{cases} \frac{1}{|V|} \frac{e^{-h_i(\sigma)\sigma_i}}{e^{h_i(\sigma)\sigma_i} + e^{-h_i(\sigma)\sigma_i}} & \text{if } \tau = \sigma^i \\ 1 - \sum_{i \in V} P_{\sigma, \sigma^i}^G & \text{if } \sigma = \tau \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

where, as usual,  $\sigma^i$  denotes the configuration obtained by  $\sigma$  with spin flip at site  $i$ . Different single spin flip dynamics can also be defined, for instance with Metropolis rates.

On the other side we can define an alternative collective dynamics, that we will call *PCA dynamics*, in the following way. Define

$$H(\sigma, \sigma') = \sum_{i \in V} \left[ -h_i(\sigma)\sigma'_i + q(1 - \sigma_i\sigma'_i) \right] \quad (8)$$

where  $q > 0$ . The PCA dynamics is the Markov chain defined by the following transition probabilities

$$P_{\sigma, \sigma'}^{PCA} = \frac{e^{-H(\sigma, \sigma')}}{Z_{\sigma}} \quad (9)$$

where

$$Z_\sigma = \sum_{\tau} e^{-H(\sigma,\tau)} = w^{PCA}(\sigma) \quad (10)$$

It is a standard task to show that the chain is reversible with respect to the measure

$$\pi^{PCA}(\sigma) = \frac{\sum_{\tau} e^{-H(\sigma,\tau)}}{\sum_{\tau,\tau'} e^{-H(\tau,\tau')}} \equiv \frac{w^{PCA}(\sigma)}{\sum_{\tau} w^{PCA}(\tau)} = \frac{w^{PCA}(\sigma)}{Z^{PCA}} \quad (11)$$

Note that  $\pi^{PCA}$  turns out to be the marginal of the Gibbs measure on the space of pairs of configurations

$$\mu_2(\sigma, \tau) = \frac{e^{-H(\sigma,\tau)}}{Z^{PCA}}. \quad (12)$$

Due to the definition (8) the transition probabilities of this Markov chain can be written as a product of the transition probability of each component  $\sigma'_i$  of the new configuration, as usual for PCAs:

$$P_{\sigma,\sigma'}^{PCA} = \prod_{i \in V} P(\sigma'_i | \sigma), \quad (13)$$

where

$$P(\sigma'_i | \sigma) = \frac{\exp[\sigma'_i(h_i(\sigma) + q\sigma_i)]}{2 \cosh(h_i(\sigma) + q\sigma_i)}.$$

Our goal is to show that, for a suitable choice of “moderately” large  $q$ :

- the invariant measure of the PCA  $\pi^{PCA}$  is “close” to  $\pi^G$ ;
- the PCA updates at each time step a large number of spins.

In order to state precisely the results we obtain, we recall the total variation distance, or  $L_1$  distance, between  $\pi^G$  and  $\pi^{PCA}$  as

$$\|\pi^{PCA} - \pi^G\|_{TV} = \frac{1}{2} \sum_{\sigma} |\pi^{PCA}(\sigma) - \pi^G(\sigma)| \quad (14)$$

## 1.2 Results

Before stating our main result, we make some remarks on the relation between  $\pi^G$  and  $\pi^{PCA}$ . Note first of all that

$$w^{PCA}(\sigma) = \sum_{\tau} e^{\sum_{i \in V} [h_i(\sigma)\tau_i - q(1-\sigma_i\tau_i)]} = \sum_{I \subset V} e^{\sum_{i \in V} h_i(\sigma)\sigma_i - 2 \sum_{i \in I} h_i(\sigma)\sigma_i - 2q|I|} =$$



$$= w^G(\sigma) \prod_{i \in V} (1 + \delta \phi_i) \quad (15)$$

where  $\delta = e^{-2q}$  and

$$\phi_i = e^{-2 \sum_j J_{ij} \sigma_i \sigma_j}$$

We will call

$$f(\sigma) = \prod_{i \in V} (1 + \delta \phi_i). \quad (16)$$

It easily follows that

$$\pi^{PCA}(\sigma) = \pi^G(\sigma) \frac{f}{\pi^G(f)} \quad (17)$$

We also define the probability  $\tilde{\pi}$  by

$$\tilde{\pi}(\sigma) = \pi^G(\sigma) \frac{f^2}{\pi^G(f^2)} \quad (18)$$

**Theorem 1.1** *For any  $q \geq 0$  let  $\delta := e^{-2q}$ . Suppose:*

(a)  $\delta = \delta(|V|)$  is such that  $\lim_{|V| \rightarrow \infty} \delta^2 |V| = 0$ ;

(b) there exists  $\delta_0 > 0$  such that

$$\sup_V \sup_{\delta \in [0, \delta_0]} \frac{1}{|V|} \text{Var}_\pi \left[ \sum_{i \in V} \frac{\phi_i}{1 + \delta \phi_i} \right] < \infty \quad (19)$$

for  $\pi = \pi^{PCA}$  and  $\pi = \tilde{\pi}$ .

Then

$$\lim_{|V| \rightarrow \infty} \|\pi^{PCA} - \pi^G\|_{TV} = 0. \quad (20)$$

Condition (20) follows by controlling the decay of correlations of the family of functions  $\{\phi_i\}$ . This control can be achieved by using the *Dobrushin uniqueness condition*. More precisely, we obtain the following result.

**Proposition 1.2** *Assume*

$$\sup_i \sum_j \tanh(2|J_{i,j}|) < 1. \quad (21)$$

Then assumption (b) in Theorem 1.1 holds.

### 1.3 Discussion and open problems

#### 1- Low temperature

Hypothesis (21) is the Dobrushin condition for uniqueness of phase for the Gibbs measure  $\pi^G$ . It is needed due to the generality of the interaction we are considering. Since the crucial ingredient in the proof of the Theorem is the correlation decay, we expect that hypothesis (21) can be weakened if additional assumptions on the interaction are considered. This will be discussed in a forthcoming paper.

In the last part of this paper we discuss in detail the Curie Weiss model. In the uniqueness region we show that (20) holds with the condition  $\lim_{|V| \rightarrow \infty} \delta = 0$ , much weaker than condition (a) in Theorem 1.1, while  $\lim_{|V| \rightarrow \infty} \delta^2 |V| = 0$  suffices also at low temperature for a modified dynamics which is forced to select configurations with positive magnetization.

#### 2- Convergence to equilibrium of PCA

Note that when  $\delta = \frac{1}{|V|}$  the PCA dynamics is essentially equivalent to sequential Gibbs sampler, since the average number of spins that are updated in a time step is of order  $\delta|V|$ . A natural question is then to compare the speed of convergence to equilibrium of the PCA dynamics for  $\frac{1}{|V|} \ll \delta \ll \frac{1}{\sqrt{|V|}}$  vs single spin flip dynamics. This is of course a central problem in applications. A quantitative comparison of the two dynamics is beyond the purposes of this paper. Our aim is rather to understand to what extent sampling of Gibbs measures can be implemented by parallel dynamics.

#### 3- Applications

The PCA dynamics discussed in this paper has been introduced in [9] in order to study the clique problem on large graphs. In that case the general setup was more complicated since the canonical ensemble was considered. The excellent numerical results obtained for the clique problem, encouraged us to better understand the PCA dynamics. In [6] a phase transition in the case of random graphs was proved. In general the extension of the result of this paper to the canonical ensemble is an interesting problem, and is currently under investigation.

## 1.4 Further remarks on the parallelization of the algorithm

The results discussed in this paper may be interesting in the context of parallel computing. We give here some hints in this direction.

We first recall that the transition probability  $P_{\sigma, \sigma'}$  of the PCA is a product measure, whose marginals are given by

$$\begin{aligned} P(\sigma'_i = \sigma_i | \sigma) &= \frac{1}{1 + \exp(-2 \sum_j J_{ji} \sigma_i \sigma_j - 2q)} \\ P(\sigma'_i = -\sigma_i | \sigma) &= \frac{\exp(-2 \sum_j J_{ji} \sigma_i \sigma_j - 2q)}{1 + \exp(-2 \sum_j J_{ji} \sigma_i \sigma_j - 2q)}. \end{aligned}$$

The corresponding Markov Chain is realized by the following simple Monte Carlo algorithm: for each site  $i$  sample independent random numbers  $u_i$  from the uniform distribution on  $[0, 1]$  and flip the spin  $i$  if  $u_i < \frac{\exp(-2 \sum_j J_{ji} \sigma_i \sigma_j - 2q)}{1 + \exp(-2 \sum_j J_{ji} \sigma_i \sigma_j - 2q)}$ . Although this algorithm can be easily parallelized if several processors are available, it is not convenient in the (realistic) case of a number  $K$  of processors available much smaller than the number of sites  $|V|$ ; indeed, to perform a single step of this MCMC it is necessary to sample  $|V| \gg K$  independent random variables.

A more efficient version of the algorithm, compared to the standard single spin flip algorithm, can be obtained in the regime covered by Theorem 1.2, in particular when  $q$  is large. Note that whenever  $u_i > p := \frac{e^{-2q+2J}}{1+e^{-2q+2J}}$ , the spin  $\sigma_i$  is left unchanged; thus, for  $q$  large, the majority of the random extractions of  $u_i$  are ineffective. This suggests the following algorithm, that realizes the *same* transition probability of the PCA:

- sample a number  $k$  from a Binomial distribution of parameters  $|V|, p$ ;
- sample a subset  $I_k \subseteq V$  with  $k$  elements, from the uniform distribution over these subsets;
- for each  $i \in I_k$ , sample independent random numbers  $u_i$  from the uniform distribution on  $[0, 1]$  and flip the spin  $i$  if  $u_i < e^{-2(J+\sum_j J_{ji}\sigma_i\sigma_j)} \frac{1+e^{2J-2q}}{1+e^{-2\sum_j J_{ji}\sigma_i\sigma_j-2q}}$ .

If  $q$  is chosen so that  $|V|p$  is of the order of  $K$  (the number of processors), this algorithm can be efficiently parallelized, since the  $u_i$ 's can be sampled by different processors.

To compare the performances of this parallel algorithm with a single spin flip dynamics we fix some notations. Let  $F$  be the time needed to compute the quantity  $\sum_j J_{ji} \sigma_i \sigma_j$ ,

which is needed for the spin update of both algorithms, and  $S$  the time needed to sample a site of  $V$  from the uniform distribution. For the challenging case of *long range interaction*,  $F$  is of order  $|V|$ , and it is much larger than  $S$ . We want to compare the two times  $T_G$  and  $T_{PCA}$  needed to try to update at least once all the sites in  $V$  in the single spin flip and in the PCA dynamics respectively. Note that  $T_G$  is of order

$$(F + S)n \log n.$$

For the PCA, the time needed to select a random subset  $I_k$  of cardinality  $k$  is of order  $Sk$ , while the sampling of the random numbers  $u_i$ ,  $i \in I_k$ , takes only a time of order 1, since it can be performed in parallel by the  $K \sim |V|p$  processors. Thus  $T_{PCA}$  is of order

$$(F + Sk) \frac{n \log n}{k} \sim \frac{T_G}{K},$$

since  $Sk \ll F$ . We therefore obtain a speedup of order  $1/K$ .

In the short range case,  $S$  could be bigger than  $F$ , and, on the other hand, the single spin flip dynamics could be parallelized by other tricks. The comparison in this case is more delicate, and will be the subject of further investigations.

## 2 Proofs

### 2.1 Proof of Theorem 1.1

Using (15), (16) and (17), we have

$$\begin{aligned} \|\pi^{PCA} - \pi^G\|_{TV} &= \sum_{\sigma} \frac{w^G(\sigma)}{Z^G} \left| \frac{\pi^{PCA}(\sigma)}{\pi^G(\sigma)} - 1 \right| = \sum_{\sigma} \frac{w^G(\sigma)}{Z^G} \left| \frac{w^{PCA}(\sigma)}{w^G(\sigma)} \frac{Z^G}{Z^{PCA}} - 1 \right| \\ &= \sum_{\sigma} \frac{w^G(\sigma)}{Z^G} \left| \frac{f(\sigma)}{\pi^G(f)} - 1 \right| = \pi^G \left( \left| \frac{f(\sigma)}{\pi^G(f)} - 1 \right| \right) \leq \frac{(\text{var}_{\pi_G}(f))^{1/2}}{\pi_G(f)} \end{aligned} \quad (22)$$

Therefore we need an estimate on the dependence on  $\delta$  of the quantity

$$\Delta(\delta) = \frac{\pi_G(f^2)}{(\pi_G(f))^2} - 1 \quad (23)$$

More precisely, we want to show that

$$\frac{1}{|V|} \ln \pi^G(f^2) - \frac{2}{|V|} \ln \pi^G(f) = O(\delta^2) \quad (24)$$

Note first that, writing

$$f(\sigma) = \exp \left[ \sum_{i \in V} \log(1 + \delta \phi_i(\sigma)) \right] \quad (25)$$

we have

$$\frac{d}{d\delta} \log \pi^G[f] = \pi^{PCA} \left[ \sum_{i \in V} \frac{\phi_i}{1 + \delta \phi_i} \right] \quad (26)$$

and

$$\frac{d^2}{d^2\delta} \log \pi^G[f] = -\pi^{PCA} \left[ \sum_{i \in V} \left( \frac{\phi_i}{1 + \delta \phi_i} \right)^2 \right] + \text{Var}_{\pi^{PCA}} \left[ \sum_{i \in V} \frac{\phi_i}{1 + \delta \phi_i} \right], \quad (27)$$

where we have used (17). Analogously, using (18), we have

$$\frac{d}{d\delta} \log \pi^G[f^2] = 2\tilde{\pi} \left[ \sum_{i \in V} \frac{\phi_i}{1 + \delta \phi_i} \right] \quad (28)$$

and

$$\frac{d^2}{d^2\delta} \log \pi^G[f^2] = -2\tilde{\pi} \left[ \sum_{i \in V} \left( \frac{\phi_i}{1 + \delta \phi_i} \right)^2 \right] + 4\text{Var}_{\tilde{\pi}} \left[ \sum_{i \in V} \frac{\phi_i}{1 + \delta \phi_i} \right]. \quad (29)$$

The idea is to exploit this explicit results in order to control up to the second order an expansion of (24) around  $\delta = 0$ . Clearly the first order computed in  $\delta = 0$  exhibit an explicit cancellation, since for  $\delta = 0$  we have that  $\pi^G = \pi^{PCA} = \tilde{\pi}$ . In order to show (24), therefore, it is enough to prove that

$$\sup_V \sup_{\delta \in [0, \delta_0]} \frac{1}{|V|} \left( \left| \frac{d^2}{d^2\delta} \log \pi^G[f] \right| + \left| \frac{d^2}{d^2\delta} \log \pi^G[f^2] \right| \right) < +\infty,$$

which, by (27) and (29), follows from (19).

□

## 2.2 Föllmer's estimate

Let  $\pi$  be a probability of  $\{-1, 1\}^{\mathbb{Z}^d}$ , and denote by  $\pi(\sigma_i | \sigma_{\setminus i})$  its local specifications. Define, for  $i \neq j$ , the *Dobrushin coefficients*:

$$\gamma_{ij} = \sup_{\sigma} \left| \pi(\sigma_i = 1 | \sigma_{\setminus i}) - \pi(\sigma_i = 1 | \sigma_{\setminus i}^j) \right|.$$

Assume the so-called *Dobrushin uniqueness condition*:

$$\gamma := \sup_i \sum_j \gamma_{ij} < 1. \quad (30)$$

Denoting by  $\Gamma$  the matrix with elements  $\gamma_{ij}$ , under (30) the matrix

$$D := \sum_{n=0}^{+\infty} \Gamma^n$$

is well defined. For a function  $f : -1, 1^{\mathbb{Z}^d} \rightarrow \mathbb{R}$ , set

$$\rho_j(f) := \sup_{\sigma} |f(\sigma) - f(\sigma^j)|.$$

The following is the main result of the beautiful paper [5] by H. Föllmer (see also [11] for related results).

**Theorem 2.1**

$$|Cov_{\pi}(f, g)| \leq \frac{1}{4} \sum_{i,j} D_{ij} \rho_i(f) \rho_j(g). \quad (31)$$

**2.3 Proof of Proposition 1.2**

We begin by showing that, under (21), the Dobrushin Uniqueness condition hold for  $\pi^{PCA}$  and  $\tilde{\pi}$ , for  $\delta$  sufficiently small.

**Proposition 2.2** *Let  $\gamma_{i,j}$  be the Dobrushin coefficients for  $\pi^{PCA}$  (resp.  $\tilde{\pi}$ ). Then*

$$\gamma_{i,j} \leq \tanh(2|J_{i,j}|) + \frac{1}{2} \rho_j(\psi_{i,\delta}),$$

where  $\psi_{i,\delta}$  is defined by

$$2\psi_{i,\delta}(\sigma) = \log \frac{1 + \delta e^{-2h_i(\sigma)}}{1 + \delta e^{2h_i(\sigma)}} + \sum_{l \neq 0} \log \frac{1 + \delta \exp[-2J_{i,l}\sigma_l - 2\sigma_l h_{i,l}(\sigma)]}{1 + \delta \exp[2J_{i,l}\sigma_l - 2\sigma_l h_{i,l}(\sigma)]}$$

with

$$h_{i,l}(\sigma) = \sum_{j \neq l} J_{i,j} \sigma_j.$$

*Proof.* The proof consists in a rather direct and straightforward computation. We give the proof for  $\pi = \pi^{PCA}$ . The proof for  $\tilde{\pi}$  is similar. Set, for simplicity,  $i = 0$ . We

write

$$H_i := \log(1 + \delta\phi_i),$$

so that

$$\pi^{PCA}(\sigma) = \frac{1}{Z^{PCA}} \exp \left[ \sum_{i,j} J_{i,j} \sigma_i \sigma_j + \sum_i H_i(\sigma) \right].$$

Note that

$$\sum_{i,j} J_{i,j} \sigma_i \sigma_j = 2\sigma_0 h_0(\sigma) + C_1(\sigma_{\setminus 0}),$$

where, with  $C_1(\sigma_{\setminus 0})$  we denote all remaining terms which do not depend on  $\sigma_0$ . Similarly

$$H_0(\sigma) = \log \left( 1 + \delta e^{-2\sigma_0 h_0(\sigma)} \right) = \frac{1}{2} \sigma_0 \log \frac{1 + \delta e^{-2h_0(\sigma)}}{1 + \delta e^{2h_0(\sigma)}} + C_2(\sigma_{\setminus 0}),$$

and, for  $l \neq 0$ ,

$$\begin{aligned} H_l(\sigma) &= \log \left( 1 + \delta e^{-2\sigma_l \sum_j J_{l,j} \sigma_j} \right) = \log \left( 1 + \delta \exp[-2J_{0,l} \sigma_0 \sigma_l - 2\sigma_l h_{0,l}(\sigma)] \right) \\ &= \frac{1}{2} \sigma_0 \log \frac{(1 + \delta \exp[-2J_{0,l} \sigma_l - 2\sigma_l h_{0,l}(\sigma)])}{(1 + \delta \exp[2J_{0,l} \sigma_l - 2\sigma_l h_{0,l}(\sigma)])} + C_3(\sigma_{\setminus 0}). \end{aligned}$$

It follows that

$$\pi^{PCA}(\sigma_0 = 1 | \sigma_{\setminus 0}) = \frac{\exp[2h_0(\sigma) + \psi_{0,\delta}(\sigma)]}{2 \cosh(2h_0(\sigma) + \psi_{0,\delta}(\sigma))}.$$

Now, writing  $\psi_\delta$  for  $\psi_{0,\delta}$ , we have

$$2h_0(\sigma) + \psi_\delta(\sigma) - (2h_0(\sigma^j) + \psi_\delta(\sigma^j)) = 4J_{0,j} + \psi_\delta(\sigma) - \psi_\delta(\sigma^j).$$

Setting  $x := -2h_0(\sigma) - \psi_\delta(\sigma)$  and  $y := 4J_{0,j} + \psi_\delta(\sigma) - \psi_\delta(\sigma^j)$ , we have

$$\pi^{PCA}(\sigma_0 = 1 | \sigma_{\setminus 0}) - \pi^{PCA}(\sigma_0 = 1 | \sigma_{\setminus 0}^j) = \frac{e^{-x}}{2 \cosh(x)} - \frac{e^{-x-y}}{2 \cosh(x+y)} =: g_y(x).$$

Unless  $y = 0$  (which gives  $g_y \equiv 0$ ), the derivative  $(g_y^2)'$  vanishes only at  $x = -y/2$ , where  $g_y^2$  attains its absolute maximum  $\tanh^2(y/2)$ . This yields

$$\left| \pi^{PCA}(\sigma_0 = 1 | \sigma_{\setminus 0}) - \pi^{PCA}(\sigma_0 = 1 | \sigma_{\setminus 0}^j) \right| \leq \tanh \left( 2|J_{0,j}| + \frac{1}{2} |\psi_\delta(\sigma) - \psi_\delta(\sigma^j)| \right).$$

Since, for every  $a, b \geq 0$ ,  $\tanh(a + b) \leq \tanh(a) + b$ , the conclusion of the lemma follows.  $\square$

To complete the proof that the Dobrushin Uniqueness condition hold for  $\pi^{PCA}$  and  $\tilde{\pi}$ , for  $\delta$  sufficiently small, it is enough to show the following result.

**Lemma 2.3** *We have*

$$\sup_i \sum_j \rho_j(\psi_{i,\delta}) = O(\delta).$$

as  $\delta \rightarrow 0$ .

*Proof.* Set  $i = 0$  and  $\psi_{0,\delta} = \psi_\delta$ . The estimate for a generic  $i$  is similar. Ignoring an irrelevant factor 2

$$\begin{aligned} \rho_j(\psi_\delta) &\leq \rho_j \left( \log \frac{1 + \delta e^{-2h_0(\sigma)}}{1 + \delta e^{2h_0(\sigma)}} \right) + \rho_j \left( \log \frac{1 + \delta \exp[-2J_{0,j}\sigma_j - 2\sigma_j h_{0,j}(\sigma)]}{1 + \delta \exp[2J_{0,j}\sigma_j - 2\sigma_j h_{0,j}(\sigma)]} \right) \\ &\quad + \sum_{l \neq 0,j} \rho_j \left( \log \frac{1 + \delta \exp[-2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma)]}{1 + \delta \exp[2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma)]} \right). \end{aligned}$$

The main difficulty comes from the third term, and we only deal with it, i.e. we show that

$$\sum_j \sum_{l \neq 0,j} \rho_j \left( \log \frac{1 + \delta \exp[-2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma)]}{1 + \delta \exp[2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma)]} \right) = o(\delta).$$

Set

$$C_l(\sigma) := \log \frac{1 + \delta \exp[-2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma)]}{1 + \delta \exp[2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma)]} = \log \left( 1 - 2\delta \frac{\sinh(2J_{0,l}\sigma_l) \exp[-2\sigma_l h_{0,l}(\sigma)]}{1 + \delta \exp[2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma)]} \right).$$

It is not restrictive to assume that  $\delta$  is small so that

$$2\delta \frac{\sinh(2|J_{0,l}|) \exp[-2\sigma_l h_{0,l}(\sigma)]}{1 + \delta \exp[2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma)]} < \frac{1}{2}.$$

Since, on  $(-1/2, 1/2)$  the map  $x \mapsto \log(1 - x)$  has Lipschitz constant 2, we have

$$\begin{aligned} |C_l(\sigma) - C_l(\sigma^j)| &\leq 4\delta \sinh(2|J_{0,l}|) \left| \frac{e^{-2\sigma_l h_{0,l}(\sigma^j)}}{1 + \delta e^{2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma^j)}} - \frac{e^{-2\sigma_l h_{0,l}(\sigma)}}{1 + \delta e^{2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma)}} \right| \\ &\leq 4\delta \sinh(2|J_{0,l}|) \left| e^{-2\sigma_l h_{0,l}(\sigma^j)} - e^{-2\sigma_l h_{0,l}(\sigma)} \right| \\ &\leq 4\delta e^{2J} \sinh(2|J_{0,l}|) |h_{0,l}(\sigma^j) - h_{0,l}(\sigma)| \leq 4\delta e^{4J} |J_{0,l}| |J_{j,l}|, \end{aligned}$$



where we have also used the facts that the map  $x \mapsto \frac{x}{1+ax}$  has Lipschitz constant 1 for  $x > 0$  and  $a > 0$ , that  $|e^x - e^y| \leq e^{\max(|x|,|y|)}|x - y|$  and that  $\sinh(2|J_{0,l}|) \leq e^{2J}|J_{0,l}|$ . It follows that

$$\sum_j \sum_{l \neq 0,j} \rho_j \left( \log \frac{1 + \delta \exp[-2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma)]}{1 + \delta \exp[2J_{0,l}\sigma_l - 2\sigma_l h_{0,l}(\sigma)]} \right) \leq 4\delta e^{4J} J^2 = O(\delta).$$

□

*Proof of Proposition 1.2.* For  $\pi = \pi^{PCA}$  or  $\pi = \tilde{\pi}$  we have, by Theorem 2.1,

$$\begin{aligned} \frac{1}{|V|} \text{Var}_\pi \left[ \sum_{i \in V} \frac{\phi_i}{1 + \delta \phi_i} \right] &= \frac{1}{|V|} \sum_{i,j \in V} \text{Cov}_\pi \left( \frac{\phi_i}{1 + \delta \phi_i}, \frac{\phi_j}{1 + \delta \phi_j} \right) \\ &\leq \frac{1}{|V|} \sum_{i,j \in V} \sum_{h,k} D_{hk} \rho_h \left( \frac{\phi_i}{1 + \delta \phi_i} \right) \rho_k \left( \frac{\phi_j}{1 + \delta \phi_j} \right). \end{aligned}$$

Since the map  $x \mapsto \frac{x}{1+\delta x}$  has Lipschitz constant 1 on  $[0, +\infty)$ , we have that

$$\rho_h \left( \frac{\phi_i}{1 + \delta \phi_i} \right) \leq \rho_h(\phi_i).$$

Moreover, it is easily seen that

$$\rho_h(\phi_i) \leq e^{2J} |J_{i,h}|.$$

Therefore, since

$$\sup_i \sum_j (\Gamma^n)_{ij} \leq \gamma^n$$

which implies

$$\sup_h \sum_k |D_{hk}| \leq \frac{1}{1 - \gamma},$$

we get

$$\begin{aligned} \frac{1}{|V|} \text{Var}_\pi \left[ \sum_{i \in V} \frac{\phi_i}{1 + \delta \phi_i} \right] &\leq e^{4J} \frac{1}{|V|} \sum_{i,j \in V} \sum_{h,k \in V} D_{hk} |J_{i,h}| |J_{j,k}| \\ &\leq J^2 e^{4J} \frac{1}{|V|} \sum_{h,k} D_{hk} \leq \frac{1}{1 - \gamma} J^2 e^{4J}, \end{aligned}$$

which completes the proof of Proposition 1.2

□

**Remark 2.4** *The control of the total variation distance between  $\pi^G$  and  $\pi^{PCA}$  by*

$$\sup_{\delta \in [0, \delta_0]} \frac{1}{|V|} \text{Var}_{\pi^{PCA}} \left[ \sum_{i \in V} \frac{\phi_i}{1 + \delta \phi_i} \right], \quad (32)$$

*may be useful even if the above quantity diverges as  $|V| \rightarrow +\infty$ . For instance, for the one dimensional model with*

$$J_{i,j} := \frac{J_1}{|i-j|^2},$$

*it is known the existence of an intermediate phase for which the spin-spin correlations decay as  $|i-j|^{-2+\varepsilon}$ , for some  $\varepsilon \in (0, 2]$  (see [8]). Assuming that a similar decay hold for the correlations of the  $\varphi_i$ 's, the quantity in (32) is expected to behave as  $|V|^\varepsilon$  as  $|V| \uparrow +\infty$ . Thus, for Theorem 1.2 to hold, we need*

$$\lim_{|V| \uparrow +\infty} \delta^2 |V|^{1+\varepsilon} = 0$$

*that, for  $\varepsilon < 1$ , still allow  $\delta|V| \rightarrow +\infty$ , i.e. a large number of spin updates per step.*

### 3 The mean field Ising model

As an example we discuss the performances of the PCA dynamics for the *mean-field Ising model*, or Curie-Weiss model. In this section some computations are given only at a heuristic level. The rigorous treatment would be straightforward but rather lengthy.

#### 3.1 Distance between $\pi^{PCA}$ and $\pi^G$ and comparison of phase diagrams

We consider now the following mean field hamiltonian on  $\mathcal{X} := \{-1, +1\}^{\{1, \dots, n\}}$

$$H_{CW}(\sigma) = -\frac{J}{2n} \sum_{i,j} \sigma_i \sigma_j \quad (33)$$

and the corresponding pair hamiltonian

$$H_{CW}(\sigma, \sigma') = -\frac{J}{2n} \sum_{i,j} \sigma_i \sigma'_j + q \sum_i (1 - \sigma_i \sigma'_i) \quad (34)$$

By using the definitions  $m = m(\sigma) = \frac{1}{n} \sum_i \sigma_i$ , we can study the mean field model in the standard way. Indeed we have immediately

$$H_{CW}(\sigma) = -\frac{J}{2} nm^2$$

so that

$$\pi^G(m) := \sum_{\sigma: m(\sigma)=m} \pi^G(\sigma) = \frac{e^{nF(m)+o(n)}}{Z^G}$$

with

$$F(m) = \frac{J}{2} m^2 + I\left(\frac{1+m}{2}\right), \quad I(x) := -x \ln x - (1-x) \ln(1-x).$$

The small remainder  $o(n)$  is such that  $\frac{o(n)}{n} \rightarrow 0$  uniformly for  $|m| < 1 - \varepsilon$  for any fixed  $\varepsilon$ . The contribution of the magnetizations close to  $\pm 1$  can be shown to be negligible. Moreover

$$Z^G = e^{nF(m^*)+o(n)}$$

where  $m^* := \arg \max F(m)$  satisfies the standard condition for the Curie-Weiss model:

$$Jm^* = \frac{1}{2} \ln \frac{1+m^*}{1-m^*}$$

obtaining  $m^* = 0$  for  $J < 1$  while for  $J > 1$  the solutions  $m^*_+ = -m^*_-$  can be obtained graphically.

With an immediate computation we get for the function  $f$  defined in (16):

$$f(m) = \left(1 + \delta e^{-Jm}\right)^{n\frac{m+1}{2}} \left(1 + \delta e^{Jm}\right)^{n\frac{1-m}{2}} =: e^{ng(m,\delta)} \quad (35)$$

with

$$g(m, \delta) := \frac{a+b}{2} + m \frac{a-b}{2}$$

with

$$a = a(m, \delta) := \ln(1 + \delta e^{-Jm}), \quad b = b(m, \delta) := \ln(1 + \delta e^{Jm})$$

The function  $g(m) = g(m, \delta)$  as a function of  $m$ , is a  $C^\infty$  function with the following properties:

$$g(0) = \ln(1 + \delta), \quad g'(0) = 0, \quad g''(0) = -\delta J(1 - J)$$

$$g'(m) = -\delta m J \cosh(Jm) + \delta(J-1) \sinh(Jm) + o(\delta) \quad (36)$$

$$g(-m) = g(m).$$

By defining  $\bar{m} := \arg \max F(m) + \delta g(m)$  we have that the measure  $\pi^{PCA}$  is concentrated on the configurations with magnetization  $\bar{m}$ .

In the case  $J < 1$  it is immediate to verify that  $\bar{m} = m^* = 0$ . Recalling that

$$F(m) = F(0) - \frac{1-J}{2}m^2 + o(m^2),$$

the probability measures  $\pi^G$  and  $\pi^{PCA}$  can be estimated by Gaussian distributions centered in 0 with variances  $[n(1-J)]^{-1}$  and  $[n(1-J-\delta J(1-J))]^{-1}$  respectively. Therefore we can compute explicitly

$$\frac{f(m)}{\pi^G(f)} = e^{n(g(m)-g(0))}(1 + o_n(1)),$$

yielding, for small  $\delta$  and large  $n$ ,

$$\begin{aligned} \frac{\pi^G(f^2)}{(\pi^G(f))^2} - 1 &= \sum_m \left( \frac{e^{nF(m)} f(m)}{Z^G \pi^G(f)} \right) \left[ \frac{f(m)}{\pi^G(f)} - 1 \right] \\ &= \sum_m \pi^{PCA}(m) \left[ \frac{f(m)}{\pi^G(f)} - 1 \right] \simeq \sum_m \pi^{PCA}(m) \left[ e^{-n\delta J(1-J)m^2/2} - 1 \right] \\ &\simeq \sqrt{\frac{n(1-J-\delta J(1-J))}{2\pi}} \int dm e^{-n(1-J-\delta J(1-J))m^2/2} \left[ e^{-n\delta J(1-J)m^2/2} - 1 \right] \simeq \frac{J\delta}{2}. \end{aligned}$$

Hence we obtain in this case a result stronger than Theorem 1.2 since we do not need the hypothesis  $\delta^2 n \rightarrow 0$ , being enough that  $\delta \rightarrow 0$  as  $n \rightarrow \infty$ .

In the low temperature case  $J > 1$ , for any finite  $J$  (temperature strictly positive), again we obtain the convergence result when  $\delta^2 n \rightarrow 0$ , if the system is restricted to a single phase. From the dynamical point of view the restriction to a single phase can be obtained simply by a reflecting barrier, following for instance [13]. Consider the dynamics on  $\mathcal{X}_+ := \{\sigma : m(\sigma) \geq 0\}$  obtained by generating a candidate move  $\sigma' \in \mathcal{X}$  according to (9) with the pair hamiltonian (34), and accepting it as a new state if  $\sigma' \in \mathcal{X}_+$  and adopting  $-\sigma'$  as new state if  $\sigma' \notin \mathcal{X}_+$ , obtaining in this way for  $\sigma, \sigma' \in \mathcal{X}^+$  a transition probability

$$P^+(\sigma, \sigma') = P^{PCA}(\sigma, \sigma') + P^{PCA}(\sigma, -\sigma')$$

This dynamics on  $\mathcal{X}_+$  is reversible w.r.t. the invariant measure restricted to  $\mathcal{X}_+$ : for each  $\sigma, \sigma' \in \mathcal{X}_+$

$$\begin{aligned} \pi^{PCA}|_{\mathcal{X}_+}(\sigma)P^+(\sigma, \sigma') &= \frac{\pi^{PCA}(\sigma)}{\pi^{PCA}(\mathcal{X}_+)} \left[ P^{PCA}(\sigma, \sigma') + P^{PCA}(\sigma, -\sigma') \right] \\ &= \frac{\pi^{PCA}(\sigma')}{\pi^{PCA}(\mathcal{X}_+)} \left[ P^{PCA}(\sigma', \sigma) + P^{PCA}(-\sigma', \sigma) \right] = \pi^{PCA}|_{\mathcal{X}_+}(\sigma')P^+(\sigma', \sigma), \end{aligned}$$

where we used the fact that  $\pi^{PCA}(\sigma') = \pi^{PCA}(-\sigma')$ . Moreover with the same argument used in the high temperature case, we obtain  $|m^* - \bar{m}| = \mathcal{O}(\delta)$  and

$$\begin{aligned} \frac{\pi^G(f^2)}{(\pi^G(f))^2} - 1 &= \sum_m \left( \frac{e^{nF(m)} f(m)}{Z^G \pi^G(f)} \right) \left[ \frac{f(m)}{\pi^G(f)} - 1 \right] \simeq \sum_m \pi^{PCA}(m) \left[ e^{n(g(m) - g(m^*))} - 1 \right] \\ &\sim e^{n[g(\bar{m}) - g(m^*)]} - 1 \sim e^{ng'(m^*)(\bar{m} - m^*)} - 1 = \mathcal{O}(\delta^2 n), \end{aligned}$$

where we used the fact that  $g'(m) = \mathcal{O}(\delta)$  (see (36)).

### 3.2 Convergence to equilibrium

We give an estimate of the mixing time of the PCA for high temperature ( $J < 1$ ) using a classical coupling argument. We first recall that the transition probabilities of the PCA obey the following identity:

$$P(\sigma'_i | \sigma) = \frac{e^{\sigma'_i(Jm(\sigma) + q\sigma_i)}}{2 \cosh((Jm(\sigma) + q\sigma_i))}. \quad (37)$$

Given two configurations  $\sigma^+$  and  $\sigma^-$  we will write  $\sigma^+ \succeq \sigma^-$  if  $\sigma_i^+ \geq \sigma_i^- \quad \forall i \in V$ . We will define a coupling of the transition probabilities  $P(\sigma'^+ | \sigma^+)$  and  $P(\sigma'^- | \sigma^-)$  in the following way: we extract for each site  $i \in V$  an independent random variable  $u_i$  uniformly distributed in  $[0, 1]$ . Then, using *the same* random variable for both realizations of our PCA we say that  $\sigma_i^{\pm} = -1$  if  $u_i \leq P(\sigma_i^{\pm} = -1 | \sigma^{\pm})$  and  $\sigma_i^{\pm} = +1$  otherwise. It is immediate to see that this updating rule gives a marginal distribution for both  $\sigma'^+$  and  $\sigma'^-$  which is the original distribution of our PCA chain, and that if  $\sigma^+ \succeq \sigma^-$  then also  $\sigma'^+ \succeq \sigma'^-$ . If we now denote with  $n_{\text{diff}} = \frac{n}{2}(m^+ - m^-)$  the number of sites  $i \in V$  such that  $\sigma_i^+ > \sigma_i^-$ , and with  $n'_{\text{diff}}$  the number of sites  $i \in V$  such that  $\sigma_i'^+ > \sigma_i'^-$ , using (37)

and the coupling rule we have that

$$\begin{aligned}
E(n'_{\text{diff}}|\sigma^\pm) &= \frac{n}{2}(m^+ - m^-) \left( 1 - \frac{e^{-Jm^+ - 2q}}{1 + e^{-Jm^+ - 2q}} - 1 + \frac{1}{1 + e^{Jm^- - 2q}} \right) \\
&+ \frac{n}{2}(1 + m^-) \left( \frac{e^{-Jm^- - 2q}}{1 + e^{-Jm^- - 2q}} - \frac{e^{-Jm^+ - 2q}}{1 + e^{Jm^+ - 2q}} \right) \\
&+ \frac{n}{2}(1 - m^+) \left( \frac{1}{1 + e^{Jm^- - 2q}} - \frac{1}{1 + e^{Jm^+ - 2q}} \right).
\end{aligned} \tag{38}$$

This can be written, up to the first order in  $e^{-2q} = \delta$ , as

$$E(n'_{\text{diff}}|\sigma^\pm) = n_{\text{diff}} - n\delta(m^+ \cosh Jm^+ - m^- \cosh Jm^- - \sinh Jm^+ + \sinh Jm^-) + O(\delta^2) \tag{39}$$

Using now the fact that

$$\sinh Jm^\pm \leq Jm^\pm \cosh Jm^\pm$$

we have that

$$E(n'_{\text{diff}}|\sigma^\pm) \leq n_{\text{diff}} - n\delta \frac{1-J}{J} (Jm^+ \cosh Jm^+ - \sinh Jm^-) + O(\delta^2) \tag{40}$$

and by the inequality

$$Jm^+ \cosh Jm^+ - \sinh Jm^- \geq J(m^+ - m^-)$$

we obtain finally

$$E(n'_{\text{diff}}|\sigma^\pm) \leq n_{\text{diff}}(1 - 2\delta(1 - J) + O(\delta^2)) \tag{41}$$

Consider now the coupling applied to two copies of the PCA starting with spins all +1 and -1 respectively. When  $J < 1$ , (41) shows that  $E(n_{\text{diff}}(t))$  contracts exponentially in  $t$ . Denoting by  $\tau_c$  the *coalescing* time of the two chains, we have

$$P(\tau_c > t) \leq P(n_{\text{diff}}(t) \neq 0) \leq E(n_{\text{diff}}(t)) \leq 2n [1 - 2\delta(1 - J)]^t$$

By monotonicity, the total variation distance from equilibrium at time  $t$  of a PCA chain starting from an arbitrary configuration is bounded above by  $P(\tau_c > t)$ . This implies that the mixing time  $T_{\text{mix}}$  is of order  $\frac{\log n}{(1-J)\delta}$ .

**Acknowledgments:** The authors wish to thank Pierre Picco for suggesting reference [5], Errico Presutti and Alain Messager for fruitful discussions.

This work has been partially supported by the PRIN projects “Campi aleatori, percolazione ed evoluzione stocastica di sistemi con molte componenti”, “Dal microscopico al macroscopico: analisi di strutture complesse e applicazioni”, and by the GDRE 224 GREFI-MEFI-CNRS-INdAM. We also wish to thank an anonymous referee for his comments, that have suggested the remarks in section 1.4.

## References

- [1] E.N.M.CIRILLO, F.R.NARDI *Metastability for Stochastic Dynamics with a Parallel Heat Bath Updating Rule*, JSP **110**, 183-217 (2003).
- [2] E.N.M.CIRILLO, F.R.NARDI, C.SPITONI *Metastability for reversible Probabilistic Cellular Automata with self-interaction*, JSP **132**, 431-471 (2008).
- [3] D. DAWSON, A., *Synchronous and asynchronous reversible Markov systems*, Canad. Math. Bull. **17** no. 5, 633-649 (1974/75).
- [4] R. FERNANDEZ, A. TOOM, *Non Gibbsianness of the invariant measures of non-reversible cellular automata with totally asymmetric noise*, Astérisque **287**, 71-87 (2003)

- [5] H. FÖLLMER, *A Covariance Estimate for Gibbs measures*, J. Funct. Anal. **46** 387-395 (1982).
- [6] A.GAUDILLIÈRE, B.SCOPPOLA, E.SCOPPOLA, M.VIALE , *Phase transition for the cavity approach to the clique problem on random graphs.*, JSP **145**, 1127-1155 (2011).
- [7] S. GOLDSTEIN , R. KUIK , J.L. LEBOWITZ, C. MAES, *From PCAs to equilibrium systems and back.*, Comm. Math. Phys. **125**, no. 1, 71-79 (1989).
- [8] J. Z. IMBRIE1, C. M. NEWMAN, *An Intermediate Phase with Slow Decay of Correlations in One Dimensional  $1/|i - j|^2$  Percolation, Ising and Potts Models*, Commun. Math. Phys. **118**, 303-336(1988).
- [9] A.IOVANELLA, B.SCOPPOLA, E.SCOPPOLA *Some spin glass ideas applied to the clique problem*, JSP **126**, 4/5 (2007) 895-915.
- [10] O. KOZLOV, N. VASILYEV, *Reversible Markov chains with local interaction, Multi-component random systems*, Adv. Probab. Related Topics **6**, Dekker, New York, pp. 451-469 (1980).
- [11] H. Künsch, *Decay of Correlations under Dobrushin's Uniqueness Condition and its Applications*, Commun. Math. Phys. **84**, 207-222 (1982)
- [12] J.L. LEBOWITZ, C. MAES, E.R. SPEER, *Statistical mechanics of probabilistic cellular automata*, J. Statist. Phys. **59**, no. 1-2, 117-170 (1990).
- [13] D.A.LEVIN, M.J.LUCZAK, Y.PERES *Glauber dynamics for the mean-field Ising model: cut-off, critical power law and metastability*, Probab. Th. Rel. Fields **146**, 223-265 (2010).
- [14] C.M. NEWMAN, *A general central limit theorem for FKG systems*, Comm. Math. Phys. Volume **91**, Number 1, 75-80, (1983).
- [15] B. SCOPPOLA, *Exact Solution for a Class of Random Walk on the Hypercube*, JSP Volume **143** (3), 413-419 (2011).